Computers & Graphics 48 (2015) 48–59

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical Section

Enhancement of range data using a structure-aware filter

Min Ki Park [a], Yong Yi Lee [a], In Yeop Jang [b], Kwan H. Lee [a,*]

[a] School of Mechatronics, Gwangju Institute of Science and Technology (GIST), South Korea
[b] Korea Institute of Science and Technology (KIST), South Korea

ARTICLE INFO

Article history:
Received 17 September 2014
Received in revised form
20 February 2015
Accepted 21 February 2015
Available online 11 March 2015

Keywords:
Depth enhancement
Bilateral filtering
Normal smoothing
Edge preservation
Denoising

ABSTRACT

Range data acquired by affordable IR depth sensors are significantly noisy and lack shape detail, making it difficult to obtain high-quality scans using depth cameras. Most methods to enhance noisy data do not preserve the original structure of the measured scene during data refinement, particularly at sharp edges and object boundaries. We present a novel approach to enhance the range data in a structure-aware manner based on the normal information in a scene. We first computed reliable normals by smoothing out noise and sharpening the edges of individual points using color-guided normal smoothing. Based on the normal information, the noisy range data were refined by a 3D bilateral filtering technique while preserving ridges, valleys, and depth discontinuities. In a comparison, our method outperformed previous techniques in terms of noise suppression and faithful structure reconstruction.

1. Introduction

The proliferation of affordable and portable depth cameras with infrared sensing (e.g., the Microsoft Kinect or MESA SwissRanger time-of-flight sensor) has made commonplace the acquisition of 3D range data in real-time. These cameras are highly useful in a wide range of applications, including 3D reconstruction of indoor scenes, object tracking, depth video rendering, and human-computer interaction. Real-world 3D scene and shape acquisition is one of the more compelling applications of this technology and has several advantages over passive image-based shape construction techniques, such as the stereo [1] or shape-from-silhouette methods [2]. A depth sensor can obtain multi-view range data from a moving camera at a video frame rate and also works well for capturing texture-less surfaces and dark environments. Several recent efforts in 3D static-object and scene scanning using the depth camera [3–6] have produced full, visually appealing 3D models within reasonable timeframes. These methods accumulate multiple scans from slightly displaced viewpoints and merge them into noise-free surfaces. For instance, Cui et al. [5] presented a framework to improve 3D object depth scans taken while moving around objects and then aligning these objects via a super-resolution approach [7] and aligning them.

The quality of 3D scanning data is significantly dependent on the quality of the measured range data. However, since depth cameras are not intended to perform like high-fidelity 3D scanners, the acquired data suffers from noise, outliers, and low spatial resolution. Invalid pixels are also contained in the data due to occlusions, interferences, and reflections in the scenes. Thus, there is a need to develop ways to refine range data for further processing.

Most methods to enhance individual depth maps exploit high-resolution color images taken from CCD-cameras mounted on depth sensors or a built-in RGB cameras. Diebel and Thrun [8] developed a depth upsampling method based on Markov Random Field (MRF) by fusing a low-resolution depth map and its corresponding higher-resolution color image. Their method solves the minimization problem on MRF and decreases noise, thereby allowing the visualization of details. However, the method requires a long computation time. Kopf et al. [9] presented a joint bilateral upsampling that jointly uses depth and color data for image enhancement. It increases the lateral resolution of a raw depth map using a high-resolution guidance image as a prior, which quickly smooths noise. Yang et al. [10] proposed an iterative depth upsampling algorithm that applies bilateral filtering [11] to a cost volume that is built based on the distance between the potential depth candidate and each label of the cost volume. Likewise, a high-resolution color image is incorporated into the bilateral filter for better edge preservation. In their method, the final depth value is selected by taking the winner-takes-all approach to cost volume, followed by a sub-pixel estimation step. The above methods assume that depth discontinuities in a scene often co-occur with color changes in the aligned color image. Consequently, two problems occur [12]: (i) *texture copying*, where

* Corresponding author.
E-mail address: khlee@gist.ac.kr (Y.Y. Lee).

http://dx.doi.org/10.1016/j.cag.2015.02.006
0097-8493/

texture patterns in the corresponding color image appear in the depth map and (ii) *edge blurring*, where actual discontinuities in depth disappear in the depth map when the regions around the object boundaries have similar color.

To counteract these challenges, Chan et al. [12] presented a noise-aware filter for real-time depth upsampling, which prevents unwanted artifacts, such as texture copying, in the geometry. In contrast to previous color-guided methods [8–10], their method adaptively controls the weights for color and depth similarities using a blending function. Garcia et al. [13] also introduced a multi-lateral filter that favors depth discontinuities over those in the guidance image. They incorporated a credibility map computed by depth gradient into the filter to avoid edge blurring across the boundary edges. Jung [14] presented an adaptive joint trilateral filter that can complementarily enhance the sharpness of a depth map and image. In [14], the local similarity between the depth map and color image is taken into account to adaptively choose the range and depth filter kernels. These methods can adjust the influence of a guidance image to cope with the common problems of the color-guided methods. However, the methods do not eliminate weakening of sharp edges, as depth gradient alone cannot account for all the salient features in a scene.

Schall et al. [16] developed a different approach based on non-local means (NLM) [15] using self-similarity in data to preserve the local structure. They applied the method to denoise 3D point clouds. Inspired by Schall et al. [16], Huhle et al. [17,18] proposed a robust non-local denoising filter that jointly considers intra-patch similarity in range and color images. This method faithfully reconstructs strong discontinuities in depth and better preserves fine structures compared to classical techniques [8,9]. Park et al. [19] proposed an energy optimization framework to perform high-quality upsampling of depth maps while avoiding over-smoothing of depth edges. They formulated the depth upsampling problem as least squares optimization with the objective function consisting of data, smoothness, and NLM regularization terms.

As described above, previous depth enhancement methods work mainly under the depth-color co-occurrence assumption and thus adjust depth values by a linear combination of existing depth values based on color information, rather than respecting original structures of scene geometry. Therefore, they are not guaranteed to keep precise shape and orientation of underlying surfaces during depth adjustment.

In this paper, a new depth enhancement scheme is proposed where the surface normals over scenes are considered in the smoothing process. The normal information typically guides surface orientation and structure for high-quality 3D scene reconstruction. However, surface normals are first-order derivatives of surfaces and they are very sensitive to high-frequency noise in depth data. Though several methods to rectify surface normals over noisy range data have been presented [20–22], most still render blurry normals near sharp edges. To address this problem, we incorporated the corresponding color information in the normal refinement stage based on the Lambertian reflectance. The depth value at a pixel was then projected onto the local tangent plane by the 3D bilateral filter [23]. As a result, the new depth values obtained by our algorithm captured the original scene structure better than previous methods where true depth value is estimated by an average of nearby depths. We make the following technical contributions:

(i) A color-guided normal smoothing method to compute a reliable normal map from a noisy depth map.
(ii) A structure-preserving range data enhancement algorithm based on normal information through a 3D bilateral filter.

2. Overview

Fig. 1 shows the flowchart of the proposed depth enhancement algorithm. Our method takes a single depth map and its corresponding color image of static objects or indoor scenes as input and yields a noise-free structure-preserved surface. The key idea is to refine the noisy depth values in a structure-aware manner by incorporating surface normal information to the depth filter extended from the traditional bilateral filter [11]. As preprocessing, invalid pixels in the depth map (e.g., holes and outliers) are detected and rectified to directly compute surface normals over the whole depth map. Then, a normal deviation over a low-pass filtered version of the data is evaluated for feature analysis. The noisy normals are refined by our color-guided normal smoothing with adaptive parameters. In succession, the depth value of each pixel is projected onto the actual surface defined by the rectified normal field to guarantee structure preservation.

Fig. 2 illustrates a geometric interpretation of our method compared to a typical depth filter. In contrast to previous methods, we define the latent surface that is locally approximated over a pixel and its nearby pixels on normal information and attempt to project a noisy point onto the actual surface. This is why the proposed method is called a *structure-aware approach*.

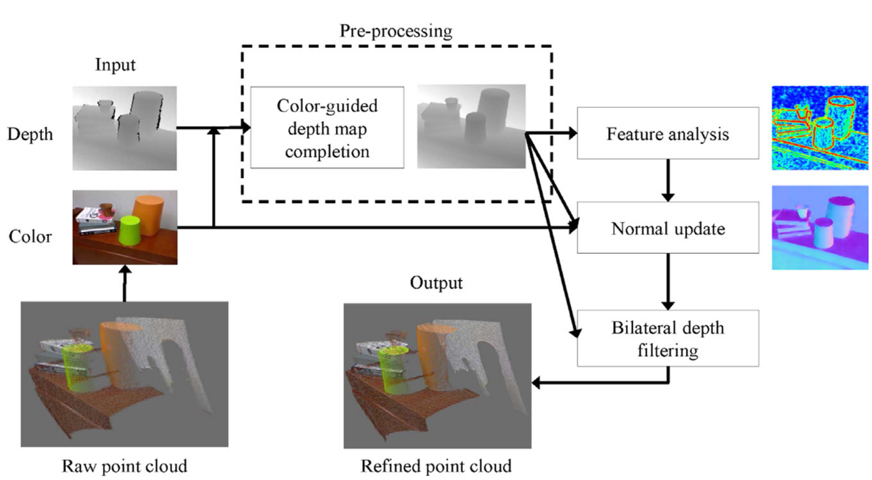

Fig. 1. Flowchart of the proposed depth enhancement algorithm.